

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/777,145

Source: _____

Date Processed by STIC: _____

ENTERED



IFWO

RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/777,145

TIME: 10:11:54

Input Set : N:\Crf3\RULE60\10777145.raw.txt

Output Set: N:\CRF4\10182004\J777145.raw

SEQUENCE LISTING

```

3 (1) GENERAL INFORMATION:
5     (i) APPLICANT: Schlessinger, Joseph
6         Sap, Jan M.
8     (ii) TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
9         PHOSPHATASE-ALPHA
11    (iii) NUMBER OF SEQUENCES: 14
13    (iv) CORRESPONDENCE ADDRESS:
14        (A) ADDRESSEE: PENNIE & EDMONDS
15        (B) STREET: 1155 AVENUE OF THE AMERICAS
16        (C) CITY: NEW YORK
17        (D) STATE: NEW YORK
18        (E) COUNTRY: U.S.A.
19        (F) ZIP: 10036
21    (v) COMPUTER READABLE FORM:
22        (A) MEDIUM TYPE: Floppy disk
23        (B) COMPUTER: IBM PC compatible
24        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
25        (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
27    (vi) CURRENT APPLICATION DATA:
C--> 28        (A) APPLICATION NUMBER: US/10/777,145
C--> 29        (B) FILING DATE: 13-Feb-2004
30        (C) CLASSIFICATION:
32    (vii) PRIOR APPLICATION DATA:
W--> 33        (A) APPLICATION NUMBER: US/09/280,597
34        (B) FILING DATE: 29-MAR-1999
W--> 35        (A) APPLICATION NUMBER: US 08/015,985
36        (B) FILING DATE: 10-FEB-1993
38    (viii) ATTORNEY/AGENT INFORMATION:
39        (A) NAME: Coruzzi, Laura A.
40        (B) REGISTRATION NUMBER: 30,742
41        (C) REFERENCE/DOCKET NUMBER: 7683-020
43    (ix) TELECOMMUNICATION INFORMATION:
44        (A) TELEPHONE: (212) 790-9090
45        (B) TELEFAX: (212) 869-9741/8864
46        (C) TELEX: 66141 PENNIE
50 (2) INFORMATION FOR SEQ ID NO: 1:
52     (i) SEQUENCE CHARACTERISTICS:
53         (A) LENGTH: 802 amino acids
54         (B) TYPE: amino acid
55         (D) TOPOLOGY: linear
57     (ii) MOLECULE TYPE: protein
59     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```

RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/777,145

TIME: 10:11:54

Input Set : N:\Crf3\RULE60\10777145.raw.txt

Output Set: N:\CRF4\10182004\J777145.raw

```

61 Met Asp Ser Trp Phe Ile Leu Val Leu Leu Gly Ser Gly Leu Ile Cys
62 1 5 10 15
64 Val Ser Ala Asn Asn Ala Thr Thr Val Ala Pro Ser Val Gly Ile Thr
65 20 25 30
67 Arg Leu Ile Asn Ser Ser Thr Ala Glu Pro Val Lys Glu Glu Ala Lys
68 35 40 45
70 Thr Ser Asn Pro Thr Ser Ser Leu Thr Ser Leu Ser Val Ala Pro Thr
71 50 55 60
73 Phe Ser Pro Asn Ile Thr Leu Gly Pro Thr Tyr Leu Thr Thr Val Asn
74 65 70 75 80
76 Ser Ser Asp Ser Asp Asn Gly Thr Thr Arg Thr Ala Ser Thr Asn Ser
77 85 90 95
79 Ile Gly Ile Thr Ile Ser Pro Asn Gly Thr Trp Leu Pro Asp Asn Gln
80 100 105 110
82 Phe Thr Asp Ala Arg Thr Glu Pro Trp Glu Gly Asn Ser Ser Thr Ala
83 115 120 125
85 Ala Thr Thr Pro Glu Thr Phe Pro Pro Ser Gly Asn Ser Asp Ser Lys
86 130 135 140
88 Asp Arg Arg Asp Glu Thr Pro Ile Ile Ala Val Met Val Ala Leu Ser
89 145 150 155 160
91 Ser Leu Leu Val Ile Val Phe Ile Ile Ile Val Leu Tyr Met Leu Arg
92 165 170 175
94 Phe Lys Lys Tyr Lys Gln Ala Gly Ser His Ser Asn Ser Phe Arg Leu
95 180 185 190
97 Ser Asn Gly Arg Thr Glu Asp Val Glu Pro Gln Ser Val Pro Leu Leu
98 195 200 205
100 Ala Arg Ser Pro Ser Thr Asn Arg Lys Tyr Pro Pro Leu Pro Val Asp
101 210 215 220
103 Lys Leu Glu Glu Glu Ile Asn Arg Arg Met Ala Asp Asp Asn Lys Leu
104 225 230 235 240
106 Phe Arg Glu Glu Phe Asn Ala Leu Pro Ala Cys Pro Ile Gln Ala Thr
107 245 250 255
109 Cys Glu Ala Ala Ser Lys Glu Glu Asn Lys Glu Lys Asn Arg Tyr Val
110 260 265 270
112 Asn Ile Leu Pro Tyr Asp His Ser Arg Val His Leu Thr Pro Val Glu
113 275 280 285
115 Gly Val Pro Asp Ser Asp Tyr Ile Asn Ala Ser Phe Ile Asn Gly Tyr
116 290 295 300
118 Gln Glu Lys Asn Lys Phe Ile Ala Ala Gln Gly Pro Lys Glu Glu Thr
119 305 310 315 320
121 Val Asn Asp Phe Trp Arg Met Ile Trp Glu Gln Asn Thr Ala Thr Ile
122 325 330 335
124 Val Met Val Thr Asn Leu Lys Glu Arg Lys Glu Cys Lys Cys Ala Gln
125 340 345 350
127 Tyr Trp Pro Asp Gln Gly Cys Trp Thr Tyr Gly Asn Ile Arg Val Ser
128 355 360 365
130 Val Glu Asp Val Thr Val Leu Val Asp Tyr Thr Val Arg Lys Phe Cys
131 370 375 380
133 Ile Gln Gln Val Gly Asp Met Thr Asn Arg Lys Pro Gln Arg Leu Ile

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/777,145

DATE: 10/18/2004

TIME: 10:11:54

Input Set : N:\Crif3\RULE60\10777145.raw.txt

Output Set: N:\CRF4\10182004\J777145.raw

134	385				390				395				400			
136	Thr	Gln	Phe	His	Phe	Thr	Ser	Trp	Pro	Asp	Phe	Gly	Val	Pro	Phe	Thr
137					405					410					415	
139	Pro	Ile	Gly	Met	Leu	Lys	Phe	Leu	Lys	Lys	Val	Lys	Ala	Cys	Asn	Pro
140				420					425					430		
142	Gln	Tyr	Ala	Gly	Ala	Ile	Val	Val	His	Cys	Ser	Ala	Gly	Val	Gly	Arg
143			435					440					445			
145	Thr	Gly	Thr	Phe	Val	Val	Ile	Asp	Ala	Met	Leu	Asp	Met	Met	His	Thr
146		450					455					460				
148	Glu	Arg	Lys	Val	Asp	Val	Tyr	Gly	Phe	Val	Ser	Arg	Ile	Arg	Ala	Gln
149	465				470					475					480	
151	Arg	Cys	Gln	Met	Val	Gln	Thr	Asp	Met	Gln	Tyr	Val	Phe	Ile	Tyr	Gln
152				485					490					495		
154	Ala	Leu	Leu	Glu	His	Tyr	Leu	Tyr	Gly	Asp	Thr	Glu	Leu	Glu	Val	Thr
155			500						505					510		
157	Ser	Leu	Glu	Thr	His	Leu	Gln	Lys	Ile	Tyr	Asn	Lys	Ile	Pro	Gly	Thr
158		515					520					525				
160	Ser	Asn	Asn	Gly	Leu	Glu	Glu	Glu	Phe	Lys	Lys	Leu	Thr	Ser	Ile	Lys
161		530				535					540					
163	Ile	Gln	Asn	Asp	Lys	Met	Arg	Thr	Gly	Asn	Leu	Pro	Ala	Asn	Met	Lys
164	545				550					555					560	
166	Lys	Asn	Arg	Val	Leu	Gln	Ile	Ile	Pro	Tyr	Glu	Phe	Asn	Arg	Val	Ile
167			565						570					575		
169	Ile	Pro	Val	Lys	Arg	Gly	Glu	Glu	Asn	Thr	Asp	Tyr	Val	Asn	Ala	Ser
170			580						585				590			
172	Phe	Ile	Asp	Gly	Tyr	Arg	Gln	Lys	Asp	Ser	Tyr	Ile	Ala	Ser	Gln	Gly
173		595					600					605				
175	Pro	Leu	Leu	His	Thr	Ile	Glu	Asp	Phe	Trp	Arg	Met	Ile	Trp	Glu	Trp
176		610				615					620					
178	Lys	Ser	Cys	Ser	Ile	Val	Met	Leu	Thr	Glu	Leu	Glu	Glu	Arg	Gly	Gln
179	625				630					635					640	
181	Glu	Lys	Cys	Ala	Gln	Tyr	Trp	Pro	Ser	Asp	Gly	Leu	Val	Ser	Tyr	Gly
182				645					650					655		
184	Asp	Ile	Thr	Val	Glu	Leu	Lys	Lys	Glu	Glu	Glu	Cys	Glu	Ser	Tyr	Thr
185			660						665				670			
187	Val	Arg	Asp	Leu	Leu	Val	Thr	Asn	Thr	Arg	Glu	Asn	Lys	Ser	Arg	Gln
188		675					680					685				
190	Ile	Arg	Gln	Phe	His	Phe	His	Gly	Trp	Pro	Glu	Val	Gly	Ile	Pro	Ser
191		690				695					700					
193	Asp	Gly	Lys	Gly	Met	Ile	Ser	Ile	Ile	Ala	Ala	Val	Gln	Lys	Gln	Gln
194	705				710					715					720	
196	Gln	Gln	Ser	Gly	Asn	His	Pro	Ile	Thr	Val	His	Cys	Ser	Ala	Gly	Ala
197			725						730				735			
199	Gly	Arg	Thr	Gly	Thr	Phe	Cys	Ala	Leu	Ser	Thr	Val	Leu	Glu	Arg	Val
200			740						745				750			
202	Lys	Ala	Glu	Gly	Ile	Leu	Asp	Val	Phe	Gln	Thr	Val	Lys	Ser	Leu	Arg
203		755					760					765				
205	Leu	Gln	Arg	Pro	His	Met	Val	Gln	Thr	Leu	Glu	Gln	Tyr	Glu	Phe	Cys
206		770				775					780					

RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/777,145

TIME: 10:11:54

Input Set : N:\Crf3\RULE60\10777145.raw.txt

Output Set: N:\CRF4\10182004\J777145.raw

```

208 Tyr Lys Val Val Gln Glu Tyr Ile Asp Ala Phe Ser Asp Tyr Ala Asn
209 785                      790                      795                      800
211 Phe Lys
214 (2) INFORMATION FOR SEQ ID NO: 2:
216     (i) SEQUENCE CHARACTERISTICS:
217         (A) LENGTH: 2409 base pairs
218         (B) TYPE: nucleic acid
219         (C) STRANDEDNESS: double
220         (D) TOPOLOGY: unknown
222     (ii) MOLECULE TYPE: cDNA
224     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
226 ATGGATTCTT GGTTCATTCT TGTTCTGCTC GGCAGTGGTC TGATATGTGT CAGTGCCAAC      60
228 AATGCTACCA CAGTTGCACC TTCTGTAGGA ATTACAAGAT TAATTAACCT ATCAACGGCA      120
230 GAACCAGTTA AAGAAGAGGC CAAAACTTCA AATCCAACCT CTCACCTAAC TTCTCTTTCT      180
232 GTGGCACCAA CATTACAGCC AAATATAACT CTGGGACCCA CCTATTTAAC CACTGTCAAT      240
234 TCTTCAGACT CTGACAATGG GACCACAAGA ACAGCAAGCA CCAATTCTAT AGGCATTACA      300
236 ATTTACCAA ATGGAACGTG GCTTCCAGAT AACCAGTTCA CGGATGCCAG AACAGAACCC      360
238 TGGGAGGGGA ATTCCAGCAC CGCAGCAACC ACTCCAGAAA CTTTCCCTCC TTCAGGTAAT      420
240 TCTGACTCGA AGGACAGAAG AGATGAGACA CCAATTATTG CGGTGATGGT GGCCCTGTCC      480
242 TCTCTGCTAG TGATCGTGTT TATTATCATA GTTTTGTTACA TGTTAAGGTT TAAGAAATAC      540
244 AAGCAAGCTG GGAGCCATTC CAATTCTTTC CGCTTATCCA ACGGCCGCAC TGAGGATGTG      600
246 GAGCCCCAGA GTGTGCCACT TCTGGCCAGA TCCCCAAGCA CCAACAGGAA ATACCCACCC      660
248 CTGCCCCTGG ACAAGCTGGA AGAGGAAATT AACCAGGAGAA TGGCAGACGA CAATAAGCTC      720
250 TTCAGGGAGG AATTCAACGC TCTCCCTGCA TGTCCCTATCC AGGCCACCTG TGAGGCTGCT      780
252 TCCAAGGAGG AAAACAAGGA AAAAAATCGA TATGTAAACA TCTTGCCCTTA TGACCACTCT      840
254 AGAGTCCACC TGACACCGGT TGAAGGGGTT CCAGATTCTG ATTACATCAA TGCTTCATTC      900
256 ATCAACGGTT ACCAAGAAAA GAACAAATTC ATTGCTGCAC AAGGACCAAA AGAAGAAACG      960
258 GTGAATGATT TCTGGCGGAT GATCTGGGAA CAAAACACAG CCACCATCGT CATGGTTACC      1020
260 AACCTGAAGG AGAGAAAGGA GTGCAAGTGC GCCCAGTACT GGCCAGACCA AGGCTGCTGG      1080
262 ACCTATGGGA ATATTCGGGT GTCTGTAGAG GATGTGACTG TCCTGGTGGA CTACACAGTA      1140
264 CGGAAGTTCT GCATCCAGCA GGTGGGCGAC ATGACCAACA GAAAGCCACA GCGCCTCATC      1200
266 ACTCAGTTCC ACTTTACCAG CTGGCCAGAC TTTGGGGTGC CTTTACCCG GATCGGCATG      1260
268 CTTAAGTTCC TCAAGAAGGT GAAGGCCTGT AACCCTCAGT ATGCAGGGGC CATCGTGGTC      1320
270 CACTGCAGTG CAGGTGTAGG GCGTACAGGT ACCTTTGTCT TCATTGATGC CATGCTGGAC      1380
272 ATGATGCATA CAGAACGGAA GGTGGACGTG TATGGCTTTG TGAGCCGGAT CCGGGCACAG      1440
274 CGCTGCCAGA TGGTGCAAAC CGATATGCAG TATGTCTTCA TATACCAAGC CCTTCTGGAG      1500
276 CATTATCTCT ATGGAATAC AGAACTGGAA GTGACCTCTC TAGAAACCCA CCTGCAGAAA      1560
278 ATTTACAACA AAATCCCAGG GACCAGCAAC AATGGATTAG AGGAGGAGTT TAAGAAGTTA      1620
280 ACATCAATCA AAATCCAGAA TGACAAGATG CGGACTGGAA ACCTTCCAGC CAACATGAAG      1680
282 AAGAACCGTG TTTTACAGAT CATTCCATAT GAATCAACA GAGTGATCAT TCCAGTTAAG      1740
284 CGGGCGAAG AGAATACAGA CTATGTGAAC GCATCCTTTA TTGATGGCTA CCGGCAGAAG      1800
286 GACTCCTATA TCGCCAGCCA GGGCCCTCTT CTCCACACAA TTGAGGACTT CTGGCGAATG      1860
288 ATCTGGGAGT GGAAATCCTG CTCTATCGTG ATGCTAACAG AACTGGAGGA GAGAGGCCAG      1920
290 GAGAAGTGTG CCCAGTACTG GCCATCTGAT GGACTGGTGT CCTATGGAGA TATTACAGTG      1980
292 GAACTGAAGA AGGAGGAGGA ATGTGAGAGC TACACCGTCC GAGACCTCCT GGTACCAAC      2040
294 ACCAGGGAGA ATAAGAGCCG GCAGATCCGG CAGTTCCACT TCCATGGCTG GCCTGAAGTG      2100
296 GGCATCCCCA GTGACGGAAA GGGCATGATC AGCATCATCG CCGCCGTGCA GAAGCAGCAG      2160
298 CAGCAGTCAG GGAACCAACC CATCACGTG CACTGCAGCG CCGGGGCAGG AAGGACGGGG      2220
300 ACCTTCTGTG CCCTGAGCAC CGTCCTGGAG CGTGTGAAAG CAGAGGGGAT TTTGGATGTC      2280

```

RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/777,145

TIME: 10:11:54

Input Set : N:\Crf3\RULE60\10777145.raw.txt

Output Set: N:\CRF4\10182004\J777145.raw

```

302 TTCCAGACTG TCAAGAGCCT GCGGCTACAG AGGCCACACA TGGTCCAGAC ACTGGAACAG      2340
304 TATGAGTTCT GCTACAAGGT GGTGCAGGAG TATATTGATG CATTCTCAGA TTATGCCAAC      2400
306 TTCAAGTAA                                     2409
309 (2) INFORMATION FOR SEQ ID NO: 3:
311     (i) SEQUENCE CHARACTERISTICS:
312         (A) LENGTH: 793 amino acids
313         (B) TYPE: amino acid
314         (D) TOPOLOGY: linear
316     (ii) MOLECULE TYPE: protein
318     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
320 Met Asp Ser Trp Phe Ile Leu Val Leu Phe Gly Ser Gly Leu Ile His
321   1           5           10           15
323 Val Ser Ala Asn Asn Ala Thr Thr Val Ser Pro Ser Leu Gly Thr Thr
324           20           25           30
326 Arg Leu Ile Lys Thr Ser Thr Thr Glu Leu Ala Lys Glu Glu Asn Lys
327           35           40           45
329 Thr Ser Asn Ser Thr Ser Ser Val Ile Ser Leu Ser Val Ala Pro Thr
330           50           55           60
332 Phe Ser Pro Asn Leu Thr Leu Glu Pro Thr Tyr Val Thr Thr Val Asn
333   65           70           75           80
335 Ser Ser His Ser Asp Asn Gly Thr Arg Arg Ala Ala Ser Thr Glu Ser
336           85           90           95
338 Gly Gly Thr Thr Ile Ser Pro Asn Gly Ser Trp Leu Ile Glu Asn Gln
339           100          105          110
341 Phe Thr Asp Ala Ile Thr Glu Pro Trp Glu Gly Asn Ser Ser Thr Ala
342           115          120          125
344 Ala Thr Thr Pro Glu Thr Phe Pro Pro Ala Asp Glu Thr Pro Ile Ile
345           130          135          140
347 Ala Val Met Val Ala Leu Ser Ser Leu Leu Val Ile Val Phe Ile Ile
348   145          150          155          160
350 Ile Val Leu Tyr Met Leu Arg Phe Lys Lys Tyr Lys Gln Ala Gly Ser
351           165          170          175
353 His Ser Asn Ser Phe Arg Leu Ser Asn Gly Arg Thr Glu Asp Val Glu
354           180          185          190
356 Pro Gln Ser Val Pro Leu Leu Ala Arg Ser Pro Ser Thr Asn Arg Lys
357           195          200          205
359 Tyr Pro Pro Leu Pro Val Asp Lys Leu Glu Glu Glu Ile Asn Arg Arg
360           210          215          220
362 Met Ala Asp Asp Asn Lys Leu Phe Arg Glu Glu Phe Asn Ala Leu Pro
363   225          230          235          240
365 Ala Cys Pro Ile Gln Ala Thr Cys Glu Ala Ala Ser Lys Glu Glu Asn
366           245          250          255
368 Lys Glu Lys Asn Arg Tyr Val Asn Ile Leu Pro Tyr Asp His Ser Arg
369           260          265          270
371 Val His Leu Thr Pro Val Glu Gly Val Pro Asp Ser Asp Tyr Ile Asn
372           275          280          285
374 Ala Ser Phe Ile Asn Gly Tyr Gln Glu Lys Asn Lys Phe Ile Ala Ala
375           290          295          300
377 Gln Gly Pro Lys Glu Glu Thr Val Asn Asp Phe Trp Arg Met Ile Trp

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/777,145

DATE: 10/18/2004
TIME: 10:11:55

Input Set : N:\Crf3\RULE60\10777145.raw.txt
Output Set: N:\CRF4\10182004\J777145.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:9; Xaa Pos.8,12,21,22,24,25,27,28,30,37,39,47,57,72,77,89,94,95,99,104
Seq#:9; Xaa Pos.109,111,115,116,124,125,131,133,135,137,138,139,143,144,153
Seq#:9; Xaa Pos.155,170,174,176,179,180,181,182,183,186,205,211,212,214,215
Seq#:9; Xaa Pos.217,222,227,230,232,240,244,247
Seq#:14; Xaa Pos.10,20,21,22,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38
Seq#:14; Xaa Pos.39,40,41,42,43,55,67,70,87,95,96,97,99,100,103,105,106,112
Seq#:14; Xaa Pos.114,115,116,120,121,123,126,127,128,133,137,138,139,141
Seq#:14; Xaa Pos.143,149,151,182,186,188,195,196,197,198,199,200,201,202
Seq#:14; Xaa Pos.205,209,212,213,214,218,222,224,228,229,244,247,254,257
Seq#:14; Xaa Pos.264,265,266,267

VERIFICATION SUMMARY

DATE: 10/18/2004

PATENT APPLICATION: US/10/777,145

TIME: 10:11:55

Input Set : N:\Crf3\RULE60\10777145.raw.txt

Output Set: N:\CRF4\10182004\J777145.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:35 M:238 W: Alpha Fields not Ordered, Reordered [(A) APPLICATION NUMBER:] of (1) (vii)
L:836 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0
M:341 Repeated in SeqNo=9
L:1142 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0
M:341 Repeated in SeqNo=14